RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2012.38.118

ScienceAsia 38 (2012): **118–124**

Improved efficient approximation of concentration parameter and confidence interval for circular distribution

S.F. Hassan, A.G. Hussin*, Y.Z. Zubairi

Centre for Foundation Studies in Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia

*Corresponding author, e-mail: ghapor@um.edu.my

Received 3 Oct 2011
Accepted 6 Feb 2012

ABSTRACT: The von Mises distribution for circular data, also known as the natural circular analogue of the normal distribution on the real line, has two parameters, namely, the concentration parameter and the circular mean. The solution of the maximum likelihood estimate (MLE) for the concentration parameter, κ, however, is analytically intractable. Thus some approximations are applied instead. In this article we propose an improved efficient approximation of κ obtained from the MLE. Unlike other estimations that have been shown to be only applicable for either large or small κ, the proposed approximation is found to be suitable for all values of κ. The improved approximation is obtained by solving piecewise polynomial equations involving the ratio of modified Bessel functions. The results of the simulation studies show that the improved approximation has a small bias and is superior to the traditional ones. Furthermore, the results obtained have been used in constructing the confidence interval for the concentration parameter, κ. Based on the simulation results, it is found that the confidence interval obtained based on the circular variance population method is superior to the confidence interval based on the normal distribution as it has smaller expected length.

KEYWORDS: von Mises distribution, Bessel function, circular data

INTRODUCTION

Circular or directional data occurs in many disciplines; for example the earth sciences, meteorology, biology, physics, psychology, image analysis, medicine, astronomy, and genetic studies [1–3]. The data are measured in the form of angles or two-dimensional orientations with the position of the zero degrees being arbitrary. This means the summary statistics used for linear data such as the arithmetic mean and the standard deviation which do not have this rotational invariance cannot be used for directional data [4–7].

For directional data, the distribution that is often used to describe its physical properties is the von Mises distribution. As a continuous probability distribution, the von Mises distribution is analogous to the normal distribution for linear data and has some similar characteristics with the normal distribution. Thus the von Mises is also known as the circular normal distribution.

A circular random variable θ follows the von Mises distribution, denoted by $M(\mu_0, \kappa)$, with probability density function $g(\theta; \mu_0, \kappa) = \{2\pi I_0(\kappa)\}^{-1} \exp\{\kappa \cos(\theta - \mu_0)\}$, where μ_0 $(0 \leqslant \mu_0 < 2\pi)$ is the mean direction and κ is known as the concentration parameter. I_0 denotes the modified Bessel function of the first kind and order zero. The Bessel functions are solutions of a second-order differential equation known as Bessel's differential equation and the probability density can also be expressed as a series of Bessel functions [8].

For large κ, the distribution becomes very concentrated around the angle μ_0 with κ being a measure of the concentration. As κ increases, the distribution of θ approaches a normal distribution with mean μ_0 and variance $1/\kappa$. As κ approaches 0, the distribution tends to converge to a uniform distribution [4–6, 9].

Some of the recent work on the von Mises distribution include a restricted maximum likelihood estimator (MLE) based on the assumption of large concentration parameters and when it is known a priori that the concentration parameters are subjected to a simple order restriction [10]. Also, an iterative algorithm using fixed points to obtain the MLE for κ in the von Mises-Fisher distribution [11] is used. Recently, the generalized von Mises [12] is extended. Matlab was used to handle the computational aspect of the parameter estimation using MLE and the trigonometric method of moments. Some applications of the von Mises distribution can be found in circular regression models

and circular linear functional models[13–15].

When estimating a parameter value, the confidence interval of the parameter provides a measure of the precision. For circular data, the confidence interval of the concentration parameter of the von Mises distribution[5,6] is often used. In some studies, a bootstrap method is used to measure the convergence of the confidence interval for a preferred direction[16].

PARAMETER ESTIMATION OF THE VON MISES DISTRIBUTION

Suppose $\theta_1, \ldots, \theta_n$ is a random sample from $M(\mu_0, \kappa)$, the MLE of the mean direction, $\bar{\theta}$ is given by

$$\bar{\theta} = \begin{cases} \tan^{-1}(S/C), & S > 0, C > 0, \\ \tan^{-1}(S/C) + \pi, & C < 0, \\ \tan^{-1}(S/C) + 2\pi, & S < 0, C > 0, \end{cases}$$

where $\bar{C} = n^{-1}\sum_{i=1}^{n}\cos\theta_i$, and $\bar{S} = n^{-1}\sum_{i=1}^{n}\sin\theta_i$. The MLE for κ, denoted by $\hat{\kappa}$ is given by the solution of $A(\hat{\kappa}) = \bar{R} = (\bar{C}^2 + \bar{S}^2)^{\frac{1}{2}}$, where $\bar{R}$ is the mean resultant length and $A(\kappa) = I_1(\kappa)/I_0(\kappa)$. By considering the inverse of the Fisher information matrix of the MLE, the variance of $\hat{\kappa}$ is given by

$$\mathrm{Var}(\hat{\kappa}) = \left\{ n\left[1 - \left(\frac{A(\hat{\kappa})}{\hat{\kappa}}\right) - A^2(\hat{\kappa})\right]\right\}^{-1}.$$

The parameter estimate $\hat{\kappa} = A^{-1}(\bar{R})$ however, cannot be evaluated simply. This is due to the presence of the modified Bessel functions in the formulation. Instead, an approximation of A^{-1} is used. The approximation can be obtained using iterative procedures in which the early version[5] includes tabulation of certain values of A^{-1}. From there on, several approximation of A^{-1} have been proposed[17–21]. Some can be quite complicated in their derivation using sophisticated computer algorithms, while some are simple and easy to derive. This paper extends previous work[20] that proposed an efficient estimate for A^{-1} for only large κ. In this study, we improve the approximation by considering both small and large values of κ. The improved estimate is compared with several approximations given in the literature. Using several measurements such as mean, bias and standard error, the performance of the proposed approximation of A^{-1} is obtained.

APPROXIMATIONS FOR THE VON MISES CONCENTRATION PARAMETER

As mentioned earlier, several approximations for $A^{-1}(x)$ for all x in $(0, 1)$ can be found in the literature. In early studies, Amos[17] proved

$$\frac{x}{\frac{1}{2} + \left(x^2 + \frac{9}{4}\right)^2} < A(x) < \frac{x}{\frac{1}{2} + \left(x^2 + \frac{1}{4}\right)^{\frac{1}{2}}}, \quad (1)$$

for $x \geqslant 0$ and hence $A^{-1}(x)$ is approximately given by

$$f(x) = \frac{x}{1-x^2}\left\{\frac{1}{2} + \left[1.46(1-x^2) + \frac{1}{4}\right]^{\frac{1}{2}}\right\}. \quad (2)$$

Later on, Mardia and Zemroch[21] provided an algorithm for calculating $A^{-1}(x)$ together with tables which were obtained iteratively. Meanwhile, by using the power series for $I_0(x)$ and $I_1(x)$, Dobson[19] gave the approximation of $A^{-1}(x)$ as

$$f(x) = \begin{cases} 2x + x^3 + \dfrac{5x^5}{6}, & x < 0.65, \\ \frac{9-8x+3x^2}{8(1-x)}, & x \geqslant 0.65, \end{cases} \quad (3)$$

and has shown that the approximation gives a lower maximum relative error than the method of Amos[17]. An improved approximation for $A^{-1}(x)$ was given by Best and Fisher[18] which is

$$f(x) = \begin{cases} 2x + x^3 + \dfrac{5x^5}{6}, & x < 0.53, \\ -0.4 + 1.39x + \frac{0.43}{1-x}, & 0.53 \leqslant x < 0.85, \\ \frac{1}{x^3 - 4x2 + 3x}, & x \geqslant 0.85, \end{cases} \quad (4)$$

in which tabulated values are given in Ref. 6.

In the following section, we describe an improvement of the approximation by identifying a threshold for values of $A(\kappa)$ in which the formulation as given by Hussin and Mohamed[20] can be applied.

PROPOSED METHOD FOR CONCENTRATION PARAMETER APPROXIMATION

In this section, we will propose the new method in approximating the concentration parameter. The new method will be validated via simulation studies with tabulated values of the concentration parameter and sample sizes.

Improved approximation based on modified Bessel functions

By definition, $A(\kappa) = I_1(\kappa)/I_0(\kappa) = t$ and from the power series for $I_0(\kappa)$ and $I_1(\kappa)$, it is found that for small κ,

$$A_s(\kappa) \approx \frac{\kappa}{2}\left\{1 - \frac{1}{8}\kappa^2 + \frac{1}{48}\kappa^4 - \ldots\right\}, \quad (5)$$

Table 1 Numerical approximation of $A(\kappa)$.

κ	$A_l(\kappa)$	$A_s(\kappa)$	$\lvert A_l(\kappa) - A_s(\kappa)\rvert$
1.40	0.5335	0.5845	0.0510
1.45	0.5547	0.6012	0.0465
1.50	0.5741	0.6182	0.0441
1.55	0.5918	0.6355	0.0436
1.60	0.6082	0.6532	0.0451
1.65	0.6232	0.6716	0.0484
1.70	0.6372	0.6908	0.0537

while for large κ,

$$A_l(\kappa) \approx 1 - \frac{1}{2\kappa} - \frac{1}{8\kappa^2} - \frac{1}{8\kappa^3} - \ldots. \qquad (6)$$

In order to find κ such that $A_s(\kappa)$ and $A_l(\kappa)$ are close to each other, it is necessary that $A_s(\kappa)/A_l(\kappa) \approx 1$. In our case, we will consider the first term of $A_s(\kappa)$ and the first two terms of $A_l(\kappa)$. Thus

$$\frac{A_s(\kappa)}{A_l(\kappa)} \approx \frac{\frac{\kappa}{2}}{1 - \frac{1}{2\kappa}} \approx 1 \qquad (7)$$

or $\kappa^2 - 2\kappa + 1 \approx 0$. Hence $\kappa^2 - 2\kappa + \delta = 0$ or $\kappa = 1 \pm \sqrt{1-\delta}$ for small δ.

The above results indicate that the threshold value is in the interval $[0, 2]$. In order to find the threshold value, a simulation study is performed for various κ that lie within the interval $[0, 2]$. The values of t_s and t_l where $A_s(\kappa) = t_s$ and $A_l(\kappa) = t_l$ are obtained where the difference between $A_s(\kappa)$ and $A_l(\kappa)$ is the smallest.

From Table 1, it can be seen that $\kappa_0 = 1.55$, where $A_l(\kappa_0) = 0.5918 = t_l$ and $A_s(\kappa_0) = 0.6355 = t_s$ give the smallest value of the absolute difference of t_l and t_s. By taking the average of t_l and t_s, we obtain a threshold value of about 0.6137.

Hence we propose

$$\hat{\kappa} = \begin{cases} A_s^{-1}(t) & t < 0.6137 \\ A_l^{-1}(t) & t \geqslant 0.6137, \end{cases} \qquad (8)$$

where

$$A_s(\kappa) = \frac{\kappa}{2}\left\{1 - \frac{1}{8}\kappa^2 + \frac{1}{48}\kappa^4 + \ldots\right\} \text{ and}$$

$$A_l(\kappa) = 1 - \frac{1}{2\kappa} - \frac{1}{8\kappa^2} - \frac{1}{8\kappa^3}\ldots. \qquad (9)$$

For $t < 0.6137$,

$$\frac{\kappa}{2}\left\{1 - \frac{1}{8}\kappa^2 + \frac{1}{48}\kappa^4 + \ldots\right\} = t \text{ or}$$

$$\kappa^5 - 6\kappa^3 + 48\kappa - 96t = 0. \qquad (10)$$

The solution of the polynomial in (9) comprises of a real root and four complex roots. The solution can be obtained numerically from several mathematical packages. E.g., in R, `polyroot(c(-96*t,48,0,-6,0,1))` would give the desired solution. The real root is the estimated value for the concentration parameter. For $t \geqslant 0.6137$, we obtain

$$1 - \frac{1}{2\kappa} - \frac{1}{8\kappa^2} - \frac{1}{8\kappa^3} = t \text{ or}$$

$$(8t-8)\kappa^3 + 4\kappa^2 + \kappa + 1 = 0. \qquad (11)$$

Similarly, the real root of the cubic polynomial in (10) corresponds to the value of the concentration parameter.

Simulation study

Computer programs were written using R to carry out the simulation study to assess the efficiency of the four different methods of approximating the concentration parameter, namely, the Amos', Dobson's, Best & Fisher's ((2)-(4)), respectively, and the new proposed method (7). Circular samples of length $n = 30$, 50, and 100 were generated from von Mises distribution with mean 0 and $\kappa = 0.5$, 1.0, 1.5, 2.0, 4.0, 6.0, 8.0, and 10.0, respectively. Let s be the number of simulations and the following quantities were obtained from the simulation study:

1. Mean, $\bar{\hat{\kappa}} = 1/s \sum \hat{\kappa}_j$,
2. Absolute Relative Estimated Bias (AREB)=$\left(\left(\left|\bar{\hat{\kappa}} - \kappa\right|\right)/\kappa\right) \times 100\%$,
3. Estimated Standard Errors (SE)=$\sqrt{1/(s-1)\sum(\hat{\kappa}_j - \bar{\hat{\kappa}})^2}$,
4. Estimated Root Mean Square Errors (RMSE)=$\sqrt{1/(s-1)\sum(\hat{\kappa}_j - \kappa)^2}$.

The simulation results with $s = 5000$ for various true values of the concentration parameter and $n = 30$, 50, and 100 are shown in Tables 2, 3, and 4, respectively. The values of the mean, absolute relative estimated bias (AREB), estimated standard error (SE) and estimated root mean square error (RMSE) were computed for all the methods. When considering mean alone, Tables 2, 3, and 4 show that the estimated mean obtained using the proposed method is very close to the true mean in most of the given κ values compared with the other three methods.

To evaluate the performance indicator of the simulation studies, the value of AREB give a better measure. From the simulation results in Tables 2, 3, and 4, it is observed that the measures of AREB for the proposed method are closer to zero for most of the values of κ as compared to the other estimates. However, it can be seen that for $n = 30$ and 50,

Table 2 Simulation results for various values of parameter concentration, κ and $n = 30$.

Performance indicator	κ	$\hat{\kappa}^{\mathrm{New}}$	$\hat{\kappa}^{\mathrm{Amos}}$	$\hat{\kappa}^{\mathrm{Best\&F}}$	$\hat{\kappa}^{\mathrm{Dob}}$
Mean	0.5	0.5950	0.5525	0.5956	0.5954
	1.0	1.0568	1.0357	1.0776	1.0736
	1.5	1.4620	1.6048	1.6036	1.5982
	2.0	2.1175	2.2066	2.1300	2.1261
	4.0	4.3814	4.8218	4.4030	4.3922
	6.0	6.5931	7.1735	6.6005	6.5977
	8.0	8.8172	9.4770	8.8211	8.8197
	10.0	11.0779	11.7895	11.0804	11.0795
AREB	0.5	18.9900	10.5100	19.1300	19.0800
	1.0	5.6800	3.5700	7.7600	7.3600
	1.5	2.5400	6.9800	6.9000	6.5500
	2.0	5.8700	10.3300	6.5000	6.3000
	4.0	9.5400	20.5500	10.0700	9.8100
	6.0	9.8800	19.5600	10.0100	9.9600
	8.0	10.2100	18.4600	10.2600	10.2500
	10.0	10.7800	17.9000	10.8000	10.8000
SE	0.5	0.2568	0.2519	0.2589	0.2582
	1.0	0.2868	0.3429	0.3258	0.3219
	1.5	0.2740	0.4604	0.4099	0.4143
	2.0	0.4825	0.6081	0.5190	0.5110
	4.0	1.1833	1.3090	1.1692	1.1776
	6.0	1.7931	1.9080	1.7889	1.7906
	8.0	2.3841	2.4865	2.3820	2.3827
	10.0	3.0345	3.1253	3.0332	3.0337
RMSE	0.5	0.2728	0.2545	0.2750	0.2742
	1.0	0.2919	0.3434	0.3340	0.3294
	1.5	0.2500	0.4685	0.4186	0.4219
	2.0	0.4941	0.6379	0.5333	0.5244
	4.0	1.2239	1.5005	1.2191	1.2228
	6.0	1.8679	2.1852	1.8669	1.8674
	8.0	2.4948	2.8321	2.4944	2.4946
	10.0	3.1993	3.5474	3.1991	3.1992

Table 3 Simulation results for various values of parameter concentration, κ and $n = 50$.

Performance indicator	κ	$\hat{\kappa}^{\mathrm{New}}$	$\hat{\kappa}^{\mathrm{Amos}}$	$\hat{\kappa}^{\mathrm{Best\&F}}$	$\hat{\kappa}^{\mathrm{Dob}}$
Mean	0.5	0.5525	0.5096	0.5526	0.5526
	1.0	1.0409	1.0050	1.0513	1.0478
	1.5	1.4520	1.5496	1.5577	1.5506
	2.0	2.0594	2.1413	2.0763	2.0773
	4.0	4.1747	4.6051	4.1959	4.1856
	6.0	6.3509	6.9265	6.3582	6.3555
	8.0	8.4384	9.0931	8.4424	8.4410
	10.0	10.5931	11.3004	10.5956	10.5947
AREB	0.5	18.9900	10.5100	19.1300	19.0800
	1.0	5.6800	3.5700	7.7600	7.3600
	1.5	2.5400	6.9800	6.9000	6.5500
	2.0	5.8700	10.3300	6.5000	6.3000
	4.0	9.5400	20.5500	10.0700	9.8100
	6.0	9.8800	19.5600	10.0100	9.9600
	8.0	10.2100	18.4600	10.2600	10.2500
	10.0	10.7800	17.9000	10.8000	10.8000
SE	0.5	0.2031	0.1960	0.2033	0.2032
	1.0	0.2282	0.2572	0.2459	0.2404
	1.5	0.2144	0.3356	0.3021	0.3114
	2.0	0.3493	0.4483	0.3817	0.3730
	4.0	0.8341	0.9501	0.8208	0.8288
	6.0	1.3271	1.4410	1.3235	1.3249
	8.0	1.7397	1.8442	1.7379	1.7386
	10.0	2.1748	2.2700	2.1737	2.1741
RMSE	0.5	0.2089	0.1929	0.2091	0.2090
	1.0	0.2308	0.2544	0.2497	0.2439
	1.5	0.2030	0.3370	0.3042	0.3128
	2.0	0.3523	0.4666	0.3883	0.3801
	4.0	0.8227	1.0647	0.8170	0.8213
	6.0	1.3408	1.6389	1.3400	1.3404
	8.0	1.7614	2.0657	1.7610	1.7612
	10.0	2.2286	2.5457	2.2284	2.2285

and for very small values of κ, that is, for $\kappa \leqslant 1.0$, the approximations by Amos seem to show the smallest AREB value. Nevertheless, as the sample size increase, specifically when $n = 100$ (see Table 4), the Amos method performs better only for $\kappa = 0.5$. It can be inferred that for large values of κ, that is for $\kappa > 1.0$, the proposed method is consistently better than the other estimates with the smallest AREB when the sample size is $n \leqslant 50$. As the sample size increases to 100, the proposed method seems to give the best estimate with the inclusive value of $\kappa = 1$. Thus it can be deduced that for sample size $n \leqslant 50$ and $\kappa > 1.0$, the proposed method is the best and as the sample size increase to 100 the proposed method is even better with a bigger range of values of κ, that is, $\kappa \geqslant 1.0$.

Other measures of the performance, namely the SE and RMSE are used. From Tables 2, 3, and 4, we can see that the values of SE and RMSE for the new proposed method are generally consistent for most of the tabulated κ values. Amos estimates give the smallest SE and RMSE for small value of κ, which is for $\kappa \leqslant 1.0$, but become large as compared to other methods for $\kappa > 1.0$. Consistent with the earlier measure of AREB, it can be deduced that the Amos estimate gives the best estimate for small κ (i.e., for $\kappa \leqslant 1.0$) but performs poorly for $\kappa > 1.0$. This suggests the superiority of the new proposed method as compared to the other two methods.

Using the measures of SE and RMSE, we note that the new proposed method gives similar values to Best & Fisher's as well as Dobson's method. How-

Table 4 Simulation results for various values of parameter concentration, κ and $n = 100$.

Performance indicator	κ	$\hat{\kappa}^{\text{New}}$	$\hat{\kappa}^{\text{Amos}}$	$\hat{\kappa}^{\text{Best\&F}}$	$\hat{\kappa}^{\text{Dob}}$
Mean	0.5	0.5279	0.4848	0.5279	0.5279
	1.0	1.0146	0.9684	1.0183	1.0171
	1.5	1.5055	1.5063	1.5216	1.5072
	2.0	2.0139	2.0893	2.0337	2.0443
	4.0	4.0653	4.4907	4.0855	4.0763
	6.0	6.1628	6.7345	6.1701	6.1674
	8.0	8.2637	8.9170	8.2677	8.2663
	10.0	10.3072	11.0122	10.3097	10.3089
AREB	0.5	10.5000	1.9100	10.5100	10.5100
	1.0	4.0900	0.5000	5.1300	4.7800
	1.5	3.2000	3.3100	3.8500	3.3700
	2.0	2.9700	7.0700	3.8200	3.8700
	4.0	4.3700	15.1300	4.9000	4.6400
	6.0	5.8500	15.4400	5.9700	5.9300
	8.0	5.4800	13.6600	5.5300	5.5100
	10.0	5.9300	13.0000	5.9600	5.9500
SE	0.5	0.1472	0.1423	0.1472	0.1472
	1.0	0.1639	0.1788	0.1698	0.1671
	1.5	0.2177	0.2263	0.2045	0.2140
	2.0	0.2343	0.3040	0.2590	0.2493
	4.0	0.5945	0.7062	0.5831	0.5898
	6.0	0.8990	1.0198	0.8960	0.8971
	8.0	1.2223	1.3372	1.2207	1.2212
	10.0	1.5120	1.6191	1.5111	1.5114
RMSE	0.5	0.1482	0.1380	0.1482	0.1482
	1.0	0.1628	0.1766	0.1688	0.1661
	1.5	0.2176	0.2263	0.2052	0.2140
	2.0	0.2301	0.3118	0.2586	0.2509
	4.0	0.5584	0.7808	0.5529	0.5569
	6.0	0.8674	1.1563	0.8668	0.8670
	8.0	1.2040	1.5178	1.2038	1.2039
	10.0	1.5008	1.8085	1.5007	1.5007

ever, those measures did not elicit the superiority of the new proposed method over the other two methods.

CONFIDENCE INTERVALS FOR CONCENTRATION PARAMETER

In this section, we will propose two different approaches in constructing the confidence intervals for the concentration parameter in von Mises distribution. Method 1 (M1) is based on the circular variance for population. Method 2 (M2) is confidence intervals based on the distribution of $\hat{\kappa}$ obtained from the Fisher information matrix.

Method 1

The confidence interval (CI) of concentration parameter κ may be obtained by considering the wrapping of the normal distribution, $N(\mu, \sigma^2)$ around the circle which gives the wrapped normal $WN(\mu, A(\kappa))$, where $A(\kappa) = \exp\{-\sigma^2/2\}$ or $\sigma^2 = -2\ln(A(\kappa))$ and from the sample circular standard deviation, v is given by $v = \{-2\ln(1-V)\}^{\frac{1}{2}}$. However, $V = 1-\bar{R}$, and hence the sample circular standard deviation can be written as

$$\begin{aligned} v &= \{-2\ln(1-(1-\bar{R}))\}^{\frac{1}{2}} \\ &= \{-2\ln(\bar{R})\}^{\frac{1}{2}} \end{aligned} \tag{12}$$

where $\bar{R}$ is the mean resultant length. By using the standard result, the $100(1-\alpha)\%$ CI for the variance, σ^2 is given by

$$\frac{(n-1)v^2}{\chi^2_{n-1,\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)v^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \tag{13}$$

Using (11) and $\sigma^2 = -2\ln(A(\kappa))$, (12) can be written as

$$\frac{(n-1)v^2}{\chi^2_{n-1,\frac{\alpha}{2}}} < -2\ln(A(\kappa)) < \frac{(n-1)v^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \tag{14}$$

Alternatively, we may write

$$Y < A(\kappa) < Z, \tag{15}$$

where $Y = \exp\left(-(n-1)v^2/(2\chi^2_{n-1,1-\alpha/2})\right)$ and $Z = \exp\left(-(n-1)v^2/(2\chi^2_{n-1,\alpha/2})\right)$. Thus we may obtain the lower value, κ_L as well as the upper value, κ_U such that $\Pr(\kappa_L < \kappa < \kappa_U) = 1-\alpha$ where $\kappa_L = A^{-1}(Y)$ and $\kappa_U = A^{-1}(Z)$, respectively. The values of $A^{-1}(Y)$ and $A^{-1}(Z)$ in (14) may be estimated based on the approximations in (9) and (10), respectively.

Method 2

Another procedure for finding the confidence intervals for κ is based on the normal distribution or asymptotic covariance of parameter obtained via the Fisher information matrix for the distribution of $\hat{\kappa}$, which is

$$\hat{\kappa} \sim N\left(\kappa, \frac{1}{n(1-\frac{\bar{R}}{\hat{\kappa}}-\bar{R}^2)}\right).$$

The 95% confidence intervals is $-B+\hat{\kappa} < \kappa < B+\hat{\kappa}$, where $B = 1.96/\left[n(1-\bar{R}/\hat{\kappa}-\bar{R}^2)\right]^{\frac{1}{2}}$.

Simulation study

Simulation studies were carried out for different sample sizes, namely $n = 30$, 50, and 100 with various values of the concentration parameter, $\kappa = 0.5$, 1.0,

Table 5 Expected length and coverage probability for various values of κ for each sample size, $n = 30$, 50, and 100.

	$n = 30$		$n = 50$		$n = 100$	
κ	M1	M2	M1	M2	M1	M2
0.5	0.761	1.026	0.587	0.810	0.415	0.583
1.0	0.948	1.261	0.742	0.961	0.534	0.670
1.5	1.133	1.553	0.834	1.169	0.562	0.812
2.0	1.559	1.973	1.140	1.467	0.777	1.010
4.0	3.880	4.196	2.850	3.110	1.941	2.132
6.0	6.195	6.457	4.590	4.815	3.120	3.291
8.0	8.484	8.704	6.267	6.468	4.274	4.432
10.0	10.796	10.981	7.935	8.118	5.421	5.569

1.5, 2.0, 4.0, 6.0, 8.0, and 10.0 for the confidence level, $\alpha = 0.05$. Without loss of generality, the mean direction will be taken as 0 during the simulation study. Let s be the number of simulations and the following computation were calculated.

1. Expected Length = Upper limit − Lower limit
2. Coverage Probability = q/s, where q = number of true value falls in the CI

The simulation studies were repeated 10 000 times. Tables 5 and 6 show the expected length and coverage probability, respectively, calculated from the simulation studies for different sample size and concentration parameter. Expected length can be defined as the class size for each CI. Smaller values of expected length imply better approximation of the CI.

From Table 5, as the value of κ increases, it can be seen that the expected length for each method increases as well. It also shows that a large concentration parameter will result in a larger expected length. Apart from that, it also noted that an increase of sample size will result in a decrease in expected length. M1 gives a consistently smaller length and hence a better approximation than M2.

For further evaluation, we also consider the coverage probability for each method. The coverage probability is the actual probability that the interval contains the true concentration parameter for each method. The simulation studies have been done at 95% of confidence level. Hence a good indicator must give the coverage probability close to 0.95. This is called the nominal coverage probability or target value. From the results obtained for M1, it can be seen that the coverage probability becomes much closer to the target value as the value of concentration parameter increases. On the other hand, M2 gives consistently higher coverage probability than the target values for all values of the concentration parameter.

Table 6 Coverage probability for various values of κ for each sample size, $n = 30$, 50, and 100.

	$n = 30$		$n = 50$		$n = 100$	
κ	M1	M2	M1	M2	M1	M2
0.5	0.880	0.974	0.862	0.962	0.853	0.955
1.0	0.885	0.966	0.910	0.957	0.908	0.968
1.5	0.866	0.972	0.843	0.965	0.802	0.956
2.0	0.846	0.965	0.897	0.960	0.911	0.970
4.0	0.928	0.960	0.925	0.954	0.920	0.947
6.0	0.939	0.960	0.937	0.958	0.939	0.954
8.0	0.947	0.962	0.940	0.952	0.944	0.954
10.0	0.947	0.961	0.947	0.958	0.941	0.951

Table 7 Estimates of κ using the new proposed method and its confidence intervals.

	Method	Estimate
κ^{New}		7.346
Expected length	M1	3.839
	M2	4.003
CI	M1	(5.563, 9.402)
	M2	(5.345, 9.348)

ILLUSTRATIVE EXAMPLES

As an illustration of the proposed method, a bivariate data set was considered. The data was collected from along the Holderness Coastline, which is the Humberside Coast of the North Sea, UK in October 1994. A total of 97 measurements of wind direction using HF radar (x) and anchored buoy (y) were recorded over a period of 22.7 days. The data was fitted using the circular regression model proposed by Downs and Mardia[14] and the fitted model is

$$\hat{y}_i = 1.253 + 2 \arctan\left\{0.906 \tan\frac{1}{2}(x_i - 1.141)\right\}.$$

In this study, we estimate the concentration parameter for the circular residuals, $\theta_i = \hat{y}_i - y_i$ based on the fitted model. Table 7 shows the concentration parameter and its expected length as well as the confidence intervals for the residuals.

From Table 7, the estimated value of the concentration parameter for the residuals is high and it can be proved using the circular plot as shown in Fig. 1. A higher concentration parameter implies that the circular residuals are highly concentrated among each other as can be seen from Fig. 1 where most of the data are scattered around $(-45°, 45°)$. Only a few observations fall outside the range.

Apart from that, it can be seen that M1 gives a shorter expected length (3.839) than M2 which gives

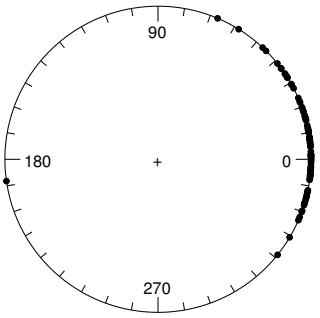

Fig. 1 Circular plot of the residuals.

4.003. This result supports the findings from the simulation results earlier, where M1 always gives the smaller expected length for all values of the concentration parameter.

REFERENCES

1. Liu D, Weinberg CR, Peddada SD (2004) A geometric approach to determine association and coherence of the activation times of cell-cycling genes under different experimental conditions. *Bioinformatics* **20**, 2521–8.
2. Liu D, Peddada SD, Li L, Weinberg CR (2006) Phase analysis of circadian-related genes in two tissues. *BMC Bioinformatics* **7**, #87.
3. Rueda C, Fernández MA, Peddada SD (2009) Estimation of parameters subject to order restriction on a circle with application to estimation of phase angles of cell-cycle genes. *J Am Stat Assoc* **104**, 338–47.
4. Mardia KV, Jupp PE (1972) *Directional Statistics*, John Wiley, Chichester.
5. Mardia KV (1972) *Statistics of Directional Data*, Academic Press, London.
6. Fisher NI (1993) *Statistical Analysis of Circular Data*. Cambridge Univ.
7. Jammalamadaka SR (2001) *Topics in Circular Statistics*. World Scientific Publishing, Singapore.
8. Abramowitz M, Stegun IG (1974) *Handbook of Mathematical Functions*. Dover Publications Inc., New York.
9. Abuzaid AH, Mohamed IB, Hussin AG (2009) A new test of discordancy in circular data. *Comm Stat Simulat Comput* **38**, 682–91.
10. Singh H, Misra N, Li S (2005) Estimation of order restricted concentration parameters of von Mises distributions. *Comm Stat Simulat Comput* **34**, 21–40.
11. Tanabe A, Fukumizu K, Oba S, Takenouchi T, Ishii S (2007) Parameter estimation for von Mises-Fisher distributions. *Comput Stat* **22**, 145–57.
12. Gatto R (2008) Some computational aspects of the generalized von Mises distribution. *J Stat Comput Simulat* **18**, 321–31.
13. Caires S, Wyatt LR (2003) A linear functional relationship model for circular data with an application to the assessment of ocean wave measurements. *J Agr Biol Environ Stat* **8**, 153–69.
14. Downs TD, Mardia KV (2002) Circular regression. *Biometrika* **89**, 683–98.
15. Hussin AG, Abuzaid AH, Zulkifli F, Mohamed IB (2010) Asymptotic covariance and detection of influential observation in a linear functional relationship model for circular data with an application to the measurements of winds directions. *Sci Asia* **36**, 249–53.
16. Otieno BS, Anderson-Coole CM (2006) On bootstrap confidence interval estimation of preferred direction for circular data. Technical paper No. 805, Grand Valley State Univ, USA.
17. Amos DE (1974) Computation of modified Bessel function and their ratios. *Math Comput* **28**, 239–42.
18. Best DJ, Fisher NI (1981) The bias of the maximum likelihood estimators of the von Mises-Fisher concentration parameters. *Comm Stat Simulat Comput* **10**, 493–502.
19. Dobson AJ (1978) Simple approximations for the von Mises concentration statistic. *Appl Stat* **27**, 345–7.
20. Hussin AG, Mohamed IB (2008) Efficient approximation for the von Mises concentration parameter. *Asian J Math Stat* **1**, 165–9.
21. Mardia KV, Zemroch PJ (1975) Algorithm AS 81: Circular statistics. *Appl Stat* **24**, 147–50.